

Analisis Butir Tes Ujian Sekolah Berstandar Nasional Matematika SMP

Sri Marlina J.Tallu

Universitas Negeri Makassar, Makassar

srimarlina.aron@gmail.com

Abstract: The purpose of this research is to reveal the item characteristics of Mathematics in National Standard School Examination (*USBN*) for junior high school, academic year of 2017-2018 at East Luwu Regency. It applied ex-post facto research design with survey approach. The data provide the stages of constructing *USBN*'s items, sample questions, test conduction, and the students' answer sheet. In line with that, data collection technique was in the form of documentation. Specifically, data analysis technique to reveal the test characteristics; rationally applied Gregory's whereas the empirical analysis determined rational validity and test reliability. While test's items were analyzed through classical test theory and item response theory. In sum, the findings show that (i) Mathematics in *USBN* has met the criteria for content and internal validity, (ii) Mathematics in *USBN* has met the elements criteria of test reliability, (iii) the empirical analysis through classical test theory for difficulty levels has placed most items as in medium category and no items in easy category. The discrimination level has functioned as in effective category which is similar to distracting items.

Keywords: *USBN* (*National Standard School Examination*), classical test theory, item response theory.

Abstrak: Penelitian ini bertujuan untuk mengetahui karakteristik butir tes Ujian Sekolah Berstandar Nasional (*USBN*) Matematika SMP yang digunakan di kabupaten Luwu Timur pada tahun ajaran 2017/2018. Pendekatan yang digunakan dalam penelitian ini adalah pendekatan survei. Penelitian ini bersifat *ex-post facto*. Data dalam penelitian ini berupa : tahapan pembuatan tes *USBN*, kisi-kisi tes *USBN*, tes *USBN* dan lembaran hasil jawaban siswa. Teknik pengumpulan data diperoleh dalam bentuk dokumentasi. Teknik analisis data mengungkap karakteristik tes ; secara rasional dilakukan menggunakan model Gregory, analisis secara empirik untuk mengetahui validitas rasional dan reliabilitas tes serta analisis butir tes berdasarkan Teori Tes Klasik dan Teori Respon Butir. Hasil penelitian menunjukkan bahwa (i) Tes *USBN* Matematika telah memenuhi kriteria validitas isi dan validitas internal, (ii) Tes *USBN* Matematika telah memenuhi unsur reliabilitas tes, (iii) Berdasarkan analisis empirik menurut Teori Klasik, untuk Tingkat Kesukaran, sebagian besar butir berada pada kategori sedang namun tidak ada butir pada kategori mudah. Daya Beda berada pada kategori baik. Kemudian Pengecoh masuk kategori berfungsi efektif.

Kata Kunci : *USBN, teori tes klasik, teori respon butir*

1. Pendahuluan

Evaluasi merupakan suatu proses penetapan nilai tentang kinerja dan hasil belajar peserta didik berdasarkan informasi yang diperoleh melalui penilaian. Dalam evaluasi pendidikan, ada empat komponen yang saling terkait dan merupakan satu kesatuan yang tidak terpisahkan. Kegiatan evaluasi harus melibatkan kegiatan penilaian, pengukuran dan tes/non tes (Mansyur, Rasyid & Suratno, 2015: 6).

Ada empat istilah yang berkaitan dengan kegiatan evaluasi pendidikan, yaitu pengukuran, pengujian, asesmen dan evaluasi. Dalam Djemary Mardapi (2016: 4) dijelaskan masing-masing, pengukuran adalah penetapan angka dengan cara yang sistematis untuk menyatakan keadaan individu atau objek. Pengujian merupakan kegiatan untuk mengetahui kemampuan yang dimiliki peserta didik. Asesmen mencakup semua cara yang digunakan untuk menilai unjuk kerja individu atau kelompok. Sedangkan, evaluasi adalah penetapan nilai atau implikasi suatu kebijakan atau putusan.

Salah satu bentuk pengukuran peserta didik terhadap hasil belajar adalah tes. Tes yang baik adalah tes yang sudah teruji baik secara rasional maupun secara empirik. Sebelum tes diberikan kepada peserta tes, ada baiknya diujikan terlebih dahulu, untuk memenuhi syarat tes dengan standar tertentu dan dapat digunakan sebagai alat ukur yang baik. Tes merupakan salah satu bentuk instrumen yang digunakan untuk melakukan pengukuran. Tes biasanya diberikan kepada peserta didik setelah mengikuti kegiatan pendidikan selama selang waktu tertentu.

Djaali & Muljono (2008: 7) menguraikan beberapa macam fungsi tes di dalam dunia pendidikan. Pertama, tes berfungsi sebagai alat untuk mengukur prestasi belajar peserta didik. Kedua, tes berfungsi sebagai motivator dalam pembelajaran. Ketiga, tes dapat berfungsi dalam upaya perbaikan kualitas pembelajaran.

Selain fungsi-fungsi tersebut, tes juga memiliki tujuan-tujuan, yaitu: (1) mengetahui tingkat kemampuan peserta didik, (2) mengukur pertumbuhan dan perkembangan peserta didik, (3) mendiagnosis kesulitan belajar peserta didik, (4) mengetahui hasil pembelajaran, (5) mengetahui hasil belajar peserta didik, (6) mengetahui pencapaian kurikulum, (7) mendorong peserta didik untuk belajar, dan (8) mendorong pendidik mengajar yang lebih baik dan peserta didik belajar lebih baik. Mardapi (2008: 68). Dengan kata lain tes memiliki peran yang sangat penting dalam dunia pendidikan.

Permendikbud Nomor 3 Tahun 2017 tentang Penilaian hasil belajar oleh satuan pendidikan dan penilaian hasil belajar oleh pemerintah Pasal (1) Ayat (5) menyatakan bahwa Ujian Sekolah Berstandar Nasional (USBN) adalah kegiatan pengukuran capaian kompetensi peserta didik yang dilakukan Satuan Pendidikan dengan mengacu pada Standar Kompetensi Lulusan untuk memperoleh pengakuan atas prestasi belajar. Istilah 'berstandar nasional' dalam USBN dimaksudkan: (1) Butir USBN disusun berdasarkan kisi-kisi yang berlaku secara nasional. Kisi-kisi dikembangkan berdasarkan Standar Isi yang diturunkan menjadi kompetensi dasar sesuai dengan kurikulum yang diterapkan; (2) Butir USBN memiliki butir dari Pusat sebanyak 20-25 persen, kecuali untuk mata pelajaran yang ditetapkan di dalam Prosedur Operasional Standar (POS) USBN; dan (3) USBN dilaksanakan dengan mengacu kepada POS USBN yang ditetapkan BSNP.

Tujuan penyelenggaraan USBN yaitu (a) untuk mengukur capaian kompetensi peserta didik yang dilakukan sekolah untuk seluruh mata pelajaran dengan mengacu pada Standar Kompetensi Lulusan untuk memperoleh pengakuan atas prestasi belajar; (b) untuk meningkatkan mutu ujian sekolah; dan (c) untuk meningkatkan kompetensi guru dalam melakukan penilaian, khususnya kemampuan pengembangan butir ujian. Manfaat dari USBN ini yaitu (a) digunakan

sebagai salah satu kriteria kelulusan; (b) penguatan Musyawarah Guru Mata Pelajaran (MGMP)/ Kelompok Kerja Guru (KKG)/ Forum Tutor dalam pengembangan dan perakitan naskah butir; dan (c) sebagai dasar pemetaan, pembinaan, dan pemberian bantuan kepada satuan pendidikan untuk pemerataan dan peningkatan mutu pendidikan. Fungsi USBN adalah sebagai penentu kelulusan peserta didik dari satuan Pendidikan.

Kegiatan menganalisis butir tes harus dilakukan oleh guru atau tim pembuat butir tes untuk meningkatkan mutu tes yang telah dibuat. Kegiatan tersebut merupakan proses pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian. Tujuan penelaahan adalah untuk mengkaji dan menelaah setiap butir tes agar diperoleh tes yang bermutu sebelum tes digunakan. Selain itu, tujuan analisis butir tes juga untuk membantu meningkatkan tes melalui revisi atau membuang butir tes yang tidak efektif, serta untuk mengetahui informasi diagnostik pada siswa apakah mereka sudah atau belum memahami materi yang diajarkan.

Manfaat analisis butir Dalam panduan analisis butir tes adalah: (1) untuk menentukan tes-tes yang cacat atau tidak berfungsi penggunaannya; (2) Untuk meningkatkan butir tes melalui tiga komponen analisis yaitu tingkat kesukaran, daya pembeda, dan pengecoh, serta meningkatkan pembelajaran melalui ambiguitas tes dan keterampilan tertentu yang menyebabkan peserta didik merasa sulit.

Dalam melakukan analisis butir-butir tes, untuk memperoleh butir tes yang berkualitas tinggi, analisis butir dapat dilakukan dengan dua pendekatan yaitu pendekatan Teori Klasik dan Teori Respon Butir. Kegiatan menganalisis butir tes selama ini jarang dilakukan. Itulah sebabnya validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan *distraktor* tes yang diujikan pada USBN perlu dipertanyakan. Apakah tes USBN yang selama ini dibuat sudah memenuhi syarat-syarat tes yang baik. Jika tidak memenuhi syarat tes yang baik, maka dapat berpengaruh terhadap kemampuan peserta didik dalam mengerjakan tes. Kelemahan ini akan berdampak pada sulitnya menentukan kemampuan peserta didik yang sebenarnya.

Karakteristik butir tes dapat terungkap dengan menggunakan dua analisis yakni analisis secara rasional untuk menentukan validitas isi dilakukan menggunakan model Gregory, dan analisis secara empirik untuk mengetahui validitas internal dan reliabilitas tes serta analisis butir tes berdasarkan Teori Tes Klasik dan Teori Respon Butir.

Bukti validitas isi suatu tes prestasi belajar dilakukan melalui dua tahap, pertama adalah kesesuaian antara indikator pada kisi-kisi dengan definisi operasional instrumen tes. Tahap kedua adalah menentukan kesesuaian antara indikator dengan item instrumen.

Validitas isi Gregory memberikan metode menentukan validitas isi menyeluruh (*overall*) berdasarkan *judgements of experts*, yaitu berupa koefisien validitas isi.

Relevansi kedua pakar secara menyeluruh merupakan validitas isi Gregory, yaitu koefisien validitas isi. Koefisien validitas isi dihitung berdasarkan rumus berikut:

Validitas isi = $\frac{D}{(A+B+C+D)}$. Penentuan nilai A, B, C dan D dapat dilihat melalui gambar berikut:

		Penilai Pakar 1	
		Relevansi Lemah Skor (1-2)	Relevansi Kuat Skor (3-4)
Penilai Pakar 2	Relevansi Lemah Skor (1-2)	A	B
	Relevansi Kuat Skor (3-4)	C	D

Keterangan:

- A = kedua pakar/ahli/validator konsisten menilai butir tersebut mempunyai relevansi yang lemah terhadap indikatornya atau tujuannya sehingga butir-butir tersebut bisa dikatakan tidak valid namun reliabel.
- B = Antara pakar/ahli/validator tidak konsisten menilai butir tersebut, ada yang menilai kuat dan ada yang menilai lemah dalam hal relevansi

terhadap indikator atau tujuannya, sehingga butir-butir tersebut bisa dikatakan tidak reliabel (perlu direvisi atau diperbaiki).

C = Antara pakar/ahli/validator tidak konsisten menilai butir tersebut, ada yang menilai kuat dan ada yang menilai lemah dalam hal relevansi terhadap indikator atau tujuannya sehingga butir-butir tersebut bisa dikatakan tidak reliabel (perlu direvisi/diperbaiki).

D = Kedua pakar/ahli/validator konsisten menilai butir tersebut mempunyai relevansi yang kuat terhadap indikator atau tujuannya, sehingga butir-butir tersebut bisa dikatakan valid dan reliabel.

Kriteria suatu instrumen yang layak digunakan jika hasil dari koefisien validitas isi memiliki relevansi yang kuat. Menurut Ruslan (2009; 19) instrumen yang mempunyai validitas isi $> 0,75$ dapat dinyatakan bahwa hasil pengukuran atau intervensi yang dilakukan oleh kedua pakar adalah valid.

Untuk memperoleh tes yang berkualitas tinggi, analisis butir dilakukan dengan dua pendekatan yaitu dengan pendekatan Teori Klasik dan Teori Respon Butir. Analisis butir secara empirik menekankan pada analisis karakteristik internal tes melalui data yang diperoleh secara empirik.

Pengukuran menghasilkan skor, baik sekor responden maupun skor butir. Mereka dapat berasal dari ujian, survei, serta dari penilaian dan pengamatan. Diharapkan skor ini dapat dipercaya. Tingkat kepercayaan skor dikenal sebagai reliabilitas skor. Batas minimal suatu tes dikatakan memiliki sifat konsisten dan stabil adalah 0,70.

Selanjutnya analisis empiris, dapat menentukan validitas internal. Validitas internal yang dimaksud pada penelitian ini yakni validitas butir, dimana validitas butir bertujuan untuk melihat seberapa jauh hasil ukur butir tes konsisten dengan hasil

ukur instrumen secara keseluruhan. Karena penelitian ini data yang diperoleh berupa skor dikotomi, maka untuk menghitung koefisien korelasi antara skor butir dengan skor total instrumen digunakan koefisien korelasi biserial (r_{bis}). Validitas Internal dibuktikan dengan cara membandingkan koefisien korelasi biserial hitung dengan nilai r tabel. Jika $r_{bis} > r$ tabel, maka butir tersebut dinyatakan valid.

Teori Klasik yakni skor amatan merupakan penjumlahan dari skor sebenarnya dengan skor kesalahan pengukuran (*error score*). Kesalahan pengukuran yang dimaksud dalam Teori Klasik merupakan kesalahan acak atau tidak sistematis. Kesalahan ini merupakan penyimpangan dari skor amatan yang diperoleh dengan skor amatan yang diharapkan.

Analisis empirik menurut pendekatan Teori Klasik meliputi tingkat kesukaran (p), daya pembeda (D), dan efektivitas pengecoh. Selain itu, dengan analisis empirik pendekatan Teori Klasik juga dapat diketahui reliabilitas tes, dan kesalahan baku pengukuran. Untuk melihat tingkat kesukaran, daya pembeda, dan efektivitas pengecoh dilakukan analisis setiap butir tes.

Tingkat kesukaran suatu butir (p) didefinisikan sebagai proporsi atau presentase subjek yang menjawab butir tes tertentu dengan benar. Analisis tingkat kesulitan soal artinya mengkaji soal-soal tes dari segi kesulitannya sehingga dapat diperoleh soal-soal mana termasuk mudah, sedang dan sukar.

Formula yang digunakan untuk mengetahui tingkat kesukaran butir yaitu:

$$P_i = \frac{\sum x_i}{Sm_i N}$$
, Di mana P_i adalah proporsi menjawab benar butir i , $\sum x_i$ merupakan banyaknya peserta tes yang menjawab benar butir i , Sm_i adalah skor maksimum, dan N merupakan jumlah peserta tes.

Daya pembeda butir merupakan kemampuan butir untuk membedakan antara peserta tes yang pandai dan peserta tes yang kurang pandai. Untuk

menentukan daya pembeda dapat menggunakan analisis indeks korelasi point biserial, indeks diskriminasi, indeks korelasi biserial, dan indeks keselarasan. Untuk analisis indeks korelasi point biserial dapat menggunakan formula

$$r_{pbis} = \left[\frac{\bar{X}_i - \bar{X}}{s_x} \right] \sqrt{\frac{p_i}{1-p_i}} \quad \text{Dimana } r_{pbis}$$

merupakan Koefisien korelasi point biserial, X_i adalah Variabel kontinu, \bar{X}_i adalah Rerata skor X untuk peserta tes yang menjawab benar butir, s_x adalah Standar deviasi dari skor X, dan p_i merupakan Proporsi peserta tes yang menjawab benar butir tersebut.

Dalam setiap tes pilihan ganda selalu digunakan alternatif jawaban yang mengandung dua unsur, yaitu jawaban tepat dan jawaban yang salah sebagai pengecoh (*distractor*). Teknik yang digunakan untuk melakukan analisis pengecoh (*distraktor*) proporsi persentase menggunakan perhitungan proporsi dengan rumus $px_i = \frac{fx_i}{M} \times 100\%$, dimana M adalah jumlah responden, px_i adalah proporsi masing-masing pilihan jawaban suatu butir tes, fx_i adalah frekuensi masing-masing pilihan jawaban suatu butir tes.

Teori Respon Butir menggunakan responsi dari para responden yang cukup banyak dan bervariasi untuk menentukan karakteristik butir. Tujuan dari Teori Respon Butir adalah untuk mendapatkan data dengan akurasi yang tinggi. Terdapat tiga syarat yang biasa digunakan dalam Teori Respon Butir yang diperlukan untuk menjamin bahwa Teori Respon Butir adalah betul seperti dengan karakteristik butir yang digunakan, yakni Unidimensi, Invariansi Parameter dan Indendpendensi lokal.

Teori Respon Butir menggunakan karakteristik butir. Setiap pemakai respon butir bebas memilih model karakteristik butir yang akan digunakannya. Bersama dengan itu muncul beberapa model teori respon diantaranya Model Logistik satu

parameter, Model logistik dua parameter dan model logistik tiga parameter.

2. Metode Penelitian

Jenis penelitian ini merupakan jenis penelitian ex-post facto dengan menggunakan dua pendekatan yaitu pendekatan Rasional dan Empiris yakni akan mengungkap karakteristik butir pada perangkat soal USBN Matematika SMP dengan menggunakan Teori Klasik dan Teori Respon Butir. Data dalam penelitian ini adalah respon peserta USBN Matematika yang diperoleh dari empat SMP di Kabupaten Luwu Timur tahun pelajaran 2017/2018 yang masing masing mewakili sekolah dengan akreditasi A, akreditasi B, yang belum diakreditasi dan sekolah swasta. Sumber data yang digunakan yaitu berupa lembar jawaban siswa yang telah didokumentasi.

Subjek dalam penelitian ini adalah seluruh peserta USBN empat SMP di Kabupaten Luwu Timur tahun pelajaran 2017/2018.

Teknik pengumpulan data yang dilakukan dalam penelitian ini adalah dengan menggunakan teknik dokumentasi, yaitu dengan mengutip respon peserta USBN Matematika SMP di Kabupaten Luwu Timur tahun pelajaran 2017/2018.

Analisis butir pada penelitian ini menggunakan dua pendekatan yaitu teori klasik dan teori respon butir. Analisis butir dengan pendekatan teori klasik menggunakan program *software* komputer yakni program TAP (*Test Analysis Program*) sedngkan analisis dengan pendekatan teori respon butir menggunakan program R i386 Vr 3.5:1

3. Hasil Penelitian dan Pembahasan

Analisis Butir Tes Ujian Sekolah Berstandar Nasional mata pelajaran Matematika tingkat SMP di Kabupaten Luwu Timur, dianalisis dengan 2 cara yakni analisis secara rasional dan analisis secara empirik. Analisis secara rasional

dilakukan untuk mengetahui validasi isi soal dari segi aspek materi, konstruksi dan bahasa. Dalam penelitian ini, untuk mengetahui validitas isi Tes USBN menggunakan uji Gregory dengan menggunakan dua pakar. Kepada pakar diberi spesifikasi pengukuran dan butir ukur. Para pakar mencocokkannya dan memberikan penilaian. Adapun analisis secara empirik dilakukan berdasarkan data empirik dimana diperoleh dari lembar jawaban siswa peserta USBN. Dua pendekatan yang dilakukan dalam analisis empirik ini, yakni secara klasik dan respon butir.

Validitas isi dapat diketahui dengan menggunakan model Gregory. Hasil validasi oleh dua validator memberikan informasi bahwa koefisien validitas isi sama dengan 1. Kriteria suatu instrumen layak digunakan jika hasil dari koefisien validitas isi lebih besar atau sama dengan 0,75 ($\geq 0,75$) hal ini sesuai dengan pendapat Ruslan (2009: 19). Berdasarkan hasil perhitungan menggunakan rumus Gregory, dihasilkan koefisien Validitas isi adalah 1, maka dapat disimpulkan bahwa Validitas isi untuk USBN Matematika SMP di Kabupaten Luwu Timur Tahun Ajaran 2017/2018 memiliki relevansi kuat.

Selanjutnya, untuk hasil analisis empirik dalam penentuan validitas internal, dapat dibuktikan dengan cara membandingkan koefisien korelasi biserial hitung dengan nilai r tabel. Jika $r_{bis} > r$ tabel, maka butir tersebut dinyatakan valid. Dari hasil perhitungan, diketahui bahwa nilai r_{tabel} untuk responden 692 yakni 0,0745 dan keseluruhan butir $r_{bis} > r_{tabel}$ jadi 30 Butir Tes USBN Matematika SMP di Luwu Timur untuk validitas internal masuk kategori valid.

Analisis tingkat koefisien reliabilitas USBN Matematika SMP di Kabupaten Luwu Timur, memberikan informasi bahwa reliabilitas Tes Ujian Sekolah Berstandar Nasional Matematika SMP sebesar 0,724, nilai koefisien reliabilitas tersebut memenuhi batas suatu

tes karena melewati batas minimal suatu sifat konsisten dan stabil yakni 0,70. Hal ini berarti Tes USBN Matematika SMP di Kabupaten Luwu Timur Tahun Ajaran 2017/2018 memenuhi unsur reliabilitas tes.

Hasil analisis butir tes USBN dengan menggunakan teori klasik, menunjukkan bahwa untuk kategori Tingkat Kesukaran Butir dari 30 Butir Tes USBN, tidak ada butir yang masuk kategori mudah, 17 butir atau 56,67% butir tes dengan tingkat kesukaran kategori sedang, sedangkan butir tes dengan tingkat kesukaran kategori sukar 13 butir atau 43,33% Butir Tes. Perbandingan antara soal kategori (mudah : sedang : sukar) adalah (0 : 17 : 13). Berdasarkan analisis, untuk daya beda, Dari 30 Butir Tes USBN, 16 atau 53,33% Butir Tes mampu membedakan kemampuan peserta tes, 6 atau 20% Butir Tes dengan kategori Kurang dan perlu direvisi supaya daya bedanya menjadi lebih baik, dan 8 atau 26,67% Butir Tes dengan kategori tidak baik yang perlu di buang atau di ganti. Indeks daya pembeda berkategori tidak baik menunjukkan bahwa butir tes tersebut tidak bisa membedakan peserta tes yang berkemampuan tinggi dan peserta tes yang berkemampuan rendah. Selanjutnya, analisis efektifitas pengecoh soal dari 30 butir tes ujian akhir sekolah, 18 atau 60 % butir tes telah berfungsi efektif, sedangkan 12 atau 40 % butir tes yang tidak berfungsi efektif. Dari banyaknya butir tes dengan pengecoh berfungsi efektif dengan yang tidak efektif menunjukkan bahwa pengecoh butir tes USBN Matematika SMP di Kabupaten Luwu Timur cukup baik.

Hasil analisis butir tes USBN dengan menggunakan teori respon butir merupakan teori yang menggunakan fungsi matematika untuk menghubungkan peluang (probabilitas) menjawab benar suatu butir dengan kemampuan (*Ability*) peserta tes. Sebelum melakukan analisis, terlebih dahulu melakukan uji persyaratan dan uji kesesuaian model.

Hambleton & Swaminathan (1985:16) menyatakan bahwa ada tiga

asumsi yang mendasari Teori Respon Butir, yaitu, unidimensi, indenpendensi lokal dan invariansi parameter. Dari hasil analisis diketahui bahwa uji unidimensi, indenpendensi lokal dan invariansi parameter terpenuhi.

Selanjutnya, untuk Uji kesesuaian (*goodness of Fit test*) yang akan dilakukan dalam analisis butir tes USBN Matematika dengan menggunakan teori respon butir, yakni dengan menggunakan software komputer yakni Program R i386 Vr 3.5.1. Berikut disajikan tabel perbandingan kecocokan model 1PL, 2PL dan 3PL.

Pada analisis tersebut, data keseluruhan (692) dianalisis secara bersama-sama. Hasil analisis tersebut memberikan informasi bahwa model 3PL lebih mampu memberikan informasi yang lebih banyak dibandingkan model 1PL dan model 2PL. Hal tersebut diketahui dengan melihat Nilai *Akaike's Information Criterion* (AIC) yang paling rendah dari ketiga model 1PL, 2PL, dan 3PL. Berikut tabel perbandingan kecocokan model 1PL, 2PL, dan 3PL

Model	AIC
Model 1PL	24252.96
Model 2PL	23899.35
Model 3PL	23735.12

Berdasarkan hasil analisis diatas, maka dalam menganalisis USBN Matematika SMP di Kabupaten Luwu Timur Tahun Ajaran 2017/2018 akan menggunakan model 3PL.

Sebelumnya telah dijelaskan bahwa pada Teori Respon Butir semakin kecil koefisien daya beda menunjukkan soal dengan daya beda tidak baik (buruk), semakin besar koefisien daya beda menunjukkan soal yang mempunyai daya beda yang baik. Nilai daya beda yang baik berkisar antara 0 sampai dengan 2. Dari analisis hasil penelitian menunjukkan bahwa ada 14 Butir Tes yang mempunyai indeks daya beda antara 0 s/d 2. Sedangkan daya beda yang kurang baik

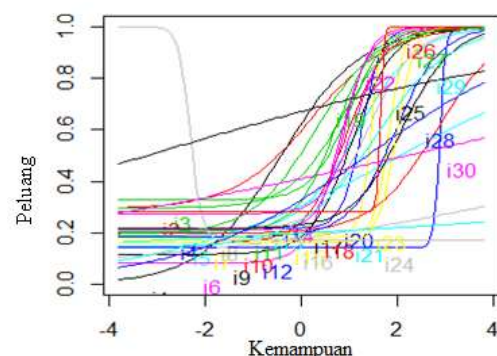
yang memiliki indeks negatif atau lebih 2 sebanyak 16 butir.

Demikian juga untuk Tingkat kesukaran Butir Tes. Indeks Tingkat Kesukaran yang baik berkisar antara -2 s/d 2. Dari hasil analisis data untuk 30 Butir Tes, sebanyak 9 Butir Tes kategori sangat mudah dan sangat sukar (indeks tingkat kesukaran berkisar < -2 atau > 2), 9 Butir Tes dengan kategori sedang. dan 12 butir tes yang memenuhi kategori sukar.

Peluang menebak merupakan ukuran peluang peserta untuk menebak dengan benar, suatu soal diharapkan mempunyai peluang menebak diharapkan kecil bahkan mendekati nol. Sehingga dikatakan soal yang baik adalah soal yang memiliki peluang menebak mendekati nol atau peluang menebak di bawah 0,25 ($c < 0,25$). Nilai peluang menebak yang kecil menunjukkan berfungsinya pengecoh dengan baik. Hasil analisis di atas dari 30 Butir Tes, sebanyak 25 Butir Tes masuk kategori peluang menebak dengan baik, dengan indeks berkisar 0,00 s/d 0,218. Sedangkan 5 Butir Tes masuk kategori peluang menebak yang tidak baik dengan indeks di atas 0,25.

Pada dasarnya untuk analisis berdasarkan teori respon butir model 3PL, butir tes baik ataupun tidak baik dapat dilihat dari bentuk grafik atau kurva karakteristiknya dan nilai ketiga parameter yang mendukung, yakni ; daya beda, tingkat kesukaran dan peluang menebak (tebakan). Berikut kurva karakteristik Butir Tes dengan model 3PL.

Kurva Karakteristik Butir



Berdasarkan kurva karakteristik 30 Butir Tes pada Gambar 4.3. merupakan bentuk kurva karakteristik Butir Tes berdasarkan model logistik 3 parameter (3PL). Butir Tes yang ideal adalah Butir Tes dengan kurva model asimpton yaitu suatu garis lengkung (linear) yang titik awal dimulai pada dasar kurva hingga titik akhir puncak kurva. Demikian juga Heri retnawati (2014) mengatakan bahwa Butir Tes yang ideal adalah Butir Tes dengan model kurva berbentuk huruf S dan cukup landai. Dengan melihat bentuk kurva karakteristik Butir Tes dan hasil analisis Butir Tes berdasarkan nilai ketiga parameter, maka dapat ditentukan butir-butir Tes yang berkualitas sesuai hasil analisis respon butir.

4. Kesimpulan

Berdasarkan pembahasan hasil penelitian di atas, maka dapat disimpulkan bahwa tes USBN Matematika telah memenuhi kriteria validitas isi, validitas internal, dan reliabilitas tes. Berdasarkan analisis empirik menurut Teori Klasik, untuk Tingkat Kesukaran, sebagian besar butir berada pada kategori sedang namun tidak ada butir pada kategori mudah. Daya Beda berada pada kategori baik. Kemudian Pengecoh masuk kategori berfungsi efektif, sedangkan menurut teori respon butir dengan menggunakan analisis model 3PL diperoleh, Tingkat Kesukaran berada pada kategori sukar, daya beda butir berada pada kategori tidak baik, sedangkan untuk tebakan semu berada pada kategori baik.

5. Daftar Pustaka

- Allen, M. J., & Yen, W. M. (1979). *Introduction to Measurement Theory*. Monterey, California: Brookd/Cole Publishing Company.
- Ali's & Khaeruddin. 2012. *Evaluasi Pembelajaran*. Makassar: Universitas Negeri Makassar
- Baso I.S. 2005. Valliditas dan Reliabilitas Tes yang Memuat Butir Dikotomi dan Politomi. *Jurnal ilmu pendidikan (Parameter) lembaga Penelitian UNJ Nomor 24 Tahun XXII*.
- Cronbach, L. J. 1984. *Essentials of Psychological Testing*. New York : Harper and Row Publisher.
- Djaali. & Muljono. 2008. *Pengukuran dalam Bidang Pendidikan*. Jakarta: Grasindo
- Depdiknas. 2010. *Panduan Analisis Butir Soal*. Jakarta: Direktorat Pembinaan SMP
- Hambleton, R.K., Swaminathan, H., 1985. *Item Response Theory*. Boston, MA: kluwer Inc.
- Mardapi, D., 2008. *Teknik Penyusunan Instrumen Tes dan Non Tes*. Yogyakarta: Mitra Cendikia
- Mardapi, Djemari. 2016. *Pengukuran, Penilaian dan Evaluasi Pendidikan*. Yogyakarta: Parama Publishing
- Mansyur, Harun & Suratno. 2015. *Asesmen Pembelajaran di Sekolah*. Makassar: Pustaka Pelajar
- Naga, Dali S. 2012. *Teori Sekor Pada Pengukuran Mental*. Jakarta: Nagarani Citrayasa
- Pakpahan, Rogers. 1999. *Studi Kebiasaan Soal Ebtanas SMU Tahun 1996/1997*. Jakarta: Jurnal kajian Dikbud No. 020 Tahun ke 5.
- Retnawati, Heri. 2009. *Analisis Butir Soal dengan Pendekatan Teori Respons Butir (online)*, (http://Jurnal_staff_uny.ac.id/files diakses 20 juli 2018)
- Ruslan, 2009. *Validitas Isi*. Buletin Pa'biritta No. 10. Tahun VI September 2009
- Whittaker, T.A, Chang W & Dodd, B.G. 2012. *The performance of IRT Model Selection Methods With Mixed – Ormat Tests*. Applied Psychological Measurement, 36(3), 159-180. Doi: 10.1177/014662161244005